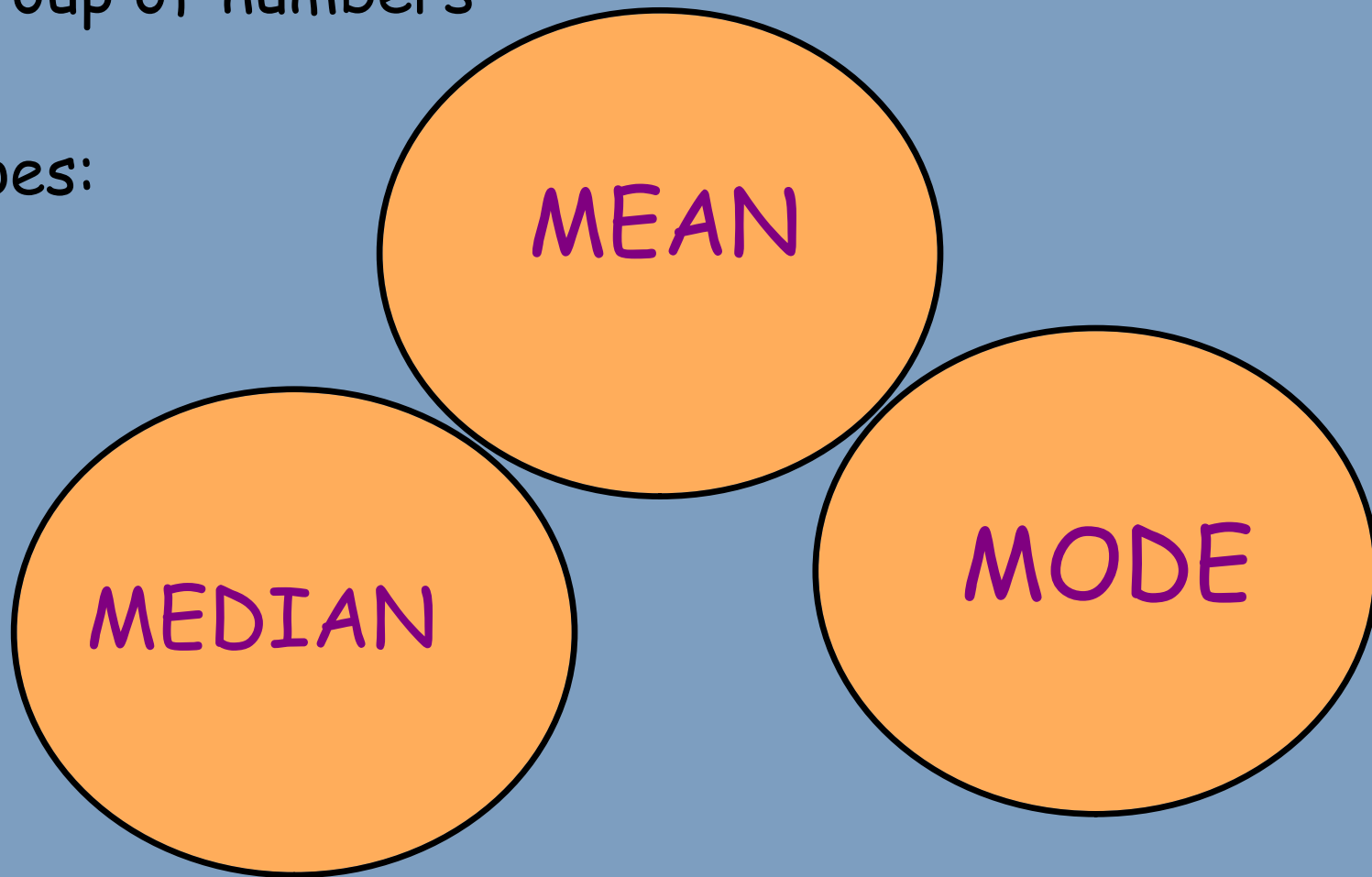


# 1.3 Describing Data

Curriculum Outcomes	Related Activities	Page in Text
<ul style="list-style-type: none"> <li>calculate various statistics using appropriate technology, analyze and interpret the displays, and describe the relationships</li> </ul>	<ul style="list-style-type: none"> <li>review computation and use of mean, median, and mode</li> </ul>	<b>15</b>
	<ul style="list-style-type: none"> <li>determine the best average to use to describe a set of reaction time measurements</li> </ul>	<b>16</b>
<ul style="list-style-type: none"> <li>calculate and apply mean and standard deviation using technology to determine if variation makes a difference</li> </ul>	<ul style="list-style-type: none"> <li>determine data values that can be treated as unusual, or outliers, and look at ways to handle them</li> </ul>	<b>15</b>
	<ul style="list-style-type: none"> <li>explore the limitations of using one piece of information to describe a set of data</li> </ul>	<b>17</b>
<ul style="list-style-type: none"> <li>create and analyze plots using appropriate technology</li> <li>make and interpret frequency bar graphs while conducting experiments and exploring measurement issues</li> </ul>	<ul style="list-style-type: none"> <li>interpret, create, and consider uses, advantages, and disadvantages of different graphs, that is, stem-and-leaf plots, box-and-whisker plots, frequency tables, histograms</li> </ul>	<b>26</b>
	<ul style="list-style-type: none"> <li>consider appropriate data grouping to create histograms</li> </ul>	<b>22</b>
<ul style="list-style-type: none"> <li>analyze statistical summaries, draw conclusions, and communicate results about distributions of data</li> </ul>	<ul style="list-style-type: none"> <li>use the shape of a graph to determine an informal measure of the spread or distribution of the data</li> </ul>	<b>22</b>

# Measures of Central Tendency

- ways to identify one number that characterizes a group of numbers
- Types:



# MODE

- the number that appears most often in a set of data.
- there can be more than one mode in a set of data.

Example:

Find the mode of this group of numbers.

20, 19, 26, 18, 26

Step 1: Arrange the numbers from least to greatest.

18, 19, 20, 26, 26

Step 2: Find the number that is repeated the most.

18, 19, 20, 26, 26

The mode is 26.

# MEAN

- the sum of a set of numbers; divided by the total number of numbers in the set. (average)

- Symbol =  $\overline{x}$

- Equation =  $\overline{x} = \frac{\text{sum of values}}{\# \text{ values}}$        $\overline{x} = \frac{\Sigma}{\# \text{ values}}$

## Example:

Find the mean of the group of numbers:

**7, 11, 12, 18, 23, 25**

Step 1: Add all the numbers.

$$7 + 11 + 12 + 18 + 23 + 25 = 96$$

Step 2: Divide the sum by the number of items.

$$96 \div 6 = 16$$

**The mean is 16.**

# MEDIAN

- the middle value or midpoint of a set of numbers arranged in order

Example:

Find the median of a group of numbers.

20, 19, 23, 18, 26

Step 1: Arrange the numbers from least to greatest.

18, 19, 20, 23, 26

Step 2: Find the middle number.

~~18~~, ~~19~~, 20, ~~23~~, ~~26~~

The median is 20.

What happens when there is 2 middle numbers??

Example:

~~1~~/8, ~~1~~/9, 20, 22, ~~2~~/3, ~~2~~/6

Step 3: Find the mean of the two middle numbers.

$$20 + 22 = 42$$

$$42 \div 2 = 21$$

The median is 21.

# Outliers

- Values that are significantly different ("lie outside") from the majority of a set of data.
- They can affect the mean.

## Example

What would the outlier of this data be?

4cm, 6cm, 5.3cm, 10.7cm, 3.2cm,  
4.6cm, 6.6cm

The outlier is 10.7cm



# Range

- the smallest number subtracted from the largest number

## Example

What would the range of this data be?

**8, 10, 6, 9, 8, 7**

**Largest = 10**

**Smallest = 6**

$$10 - 6 = 4$$

**The range is 4**

## Homework

COPY THIS DOWN

Find the mode, mean & median for each of the following:

1. 5, 12, 12, 28, 23, 31

2. 50, 67, 79, 45

3. 4, 1, 7, 3, 1, 4, 8, 9, 9

4. 25, 25, 25, 29, 27, 27

## Investigation #2

### p.14

- We are going to do this activity in pairs  
----> don't get into pairs until after we  
have gone over instructions!!

The range is 4

## Purpose

To communicate your reaction time to an event.

Investigation #2

p.14

## Procedure

- Assign one person (A) to hold a ruler straight up and down at the "30 cm" end with the "0 cm" end nearest to the floor
- The other person (B) holds their forefinger and thumb on either side of the ruler at the "0 cm" end.
- Person A drops the ruler without warning for person B to catch between the forefingers and thumb.
- Record, to the nearest millimetre, where the higher edge of the forefinger is on the ruler after the catch.
- Repeat the ruler drops twenty times.
- Switch positions so that person B now drops the ruler and drop the ruler another twenty times.
- RECORD your data in a chart in your book

Each person must make a table similar to this one in their notes

Results	
Drop Number Person A (name)	Catch Height (cm) Person A (name)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

Investigation #2  
p.14

# Questions:

Each person  
must do these  
questions with  
their 20 pieces  
of data

- Complete questions 1 & 2 on page 15
- Complete the following using the chart made for the 20 trials:

Calculate the mean, median, and range of the 20 reaction times

MEAN -> (sum of all the reaction times)/20 =

MEDIAN -> the middle value when all values are lined up least to greatest =

RANGE -> (Maximum value) - (Minimum value) =

- Does your data have any outliers? State why or why not.
- Does your data have a mode(s)?

## Warm-Up

Find the mean, median, mode, and range for each of the following:

1. 12, 4, 6, 8, 3, 6

2. 6, 1, 7, 4, 1, 4, 0, 9

## What central tendency do we use to describe our data?

- We must be careful which means of central tendency (mean, median or mode) we are going to use to find the average of a set of data.
- Not every average will give us accurate information for analyzing.

### When Mean is best:

- When the difference between the highest and the lowest piece of data is not a large difference.
- No outliers present

### Example:

The gas prices at three different stations were recorded and the following information was found:

Station #1 - \$1.24/L

Station #2 - \$1.23/L

Station #3 - \$1.25/L



### When Median is best:

- When data is clustered closely together
- Can be used when there are outliers

### Example:

There are four houses being sold in the same subdivision. They are all approximately the same age, and style. Their prices are below:

House #1: \$150,000

House #2: \$155,000

House #3: \$149,000

House #4: \$180,000

\*\* We see that there is an over priced house (an outlier) because of this we can not use the Mean as it would skew it so a Median should be used \*\*

### When Mode is best:

- rarely used
- specific to data that can be grouped

### Example:

The number of children in a family was recorded for 10 families and the following data was collected.

1, 4, 3, 4, 2, 1, 3, 2, 3, 2

\*\* Here we can group the data into families with 1 child, 2 children, 3 children, 4 children, etc. By doing this we can see what the most frequent number of children in a family are \*\*

## Questions:

Complete questions 3 and 4 on page 16.

3. a) 2, 2, 12, 13, 14, 15, 16, 17, 18 (pencil length cm)

b) 90, 100, 110, 120, 130, 140, 150, 160, 170 (duration of phone calls, s)

c) 5, 6, 6, 6, 7, 7, 8, 9, 12 (shoe sizes)

## Copy and Complete

For each of the following situations, state what the most useful average measure is and explain why. Solve for the average.

Situation 1: Nick surveyed the men in the class in regards to the size of their shoes and obtained the following data:

7, 11.5, 8, 8.5, 9, 12, 7.5, 8, 9, 7, 9

Situation 2: Rob recorded his golf scores over 7 games.

96, 92, 64, 101, 95, 93

Situation 3: Maggie compared the prices of one brand of DVD players at different stores.

\$ 409, \$399, \$429, \$439, \$399, \$429, \$419, \$389